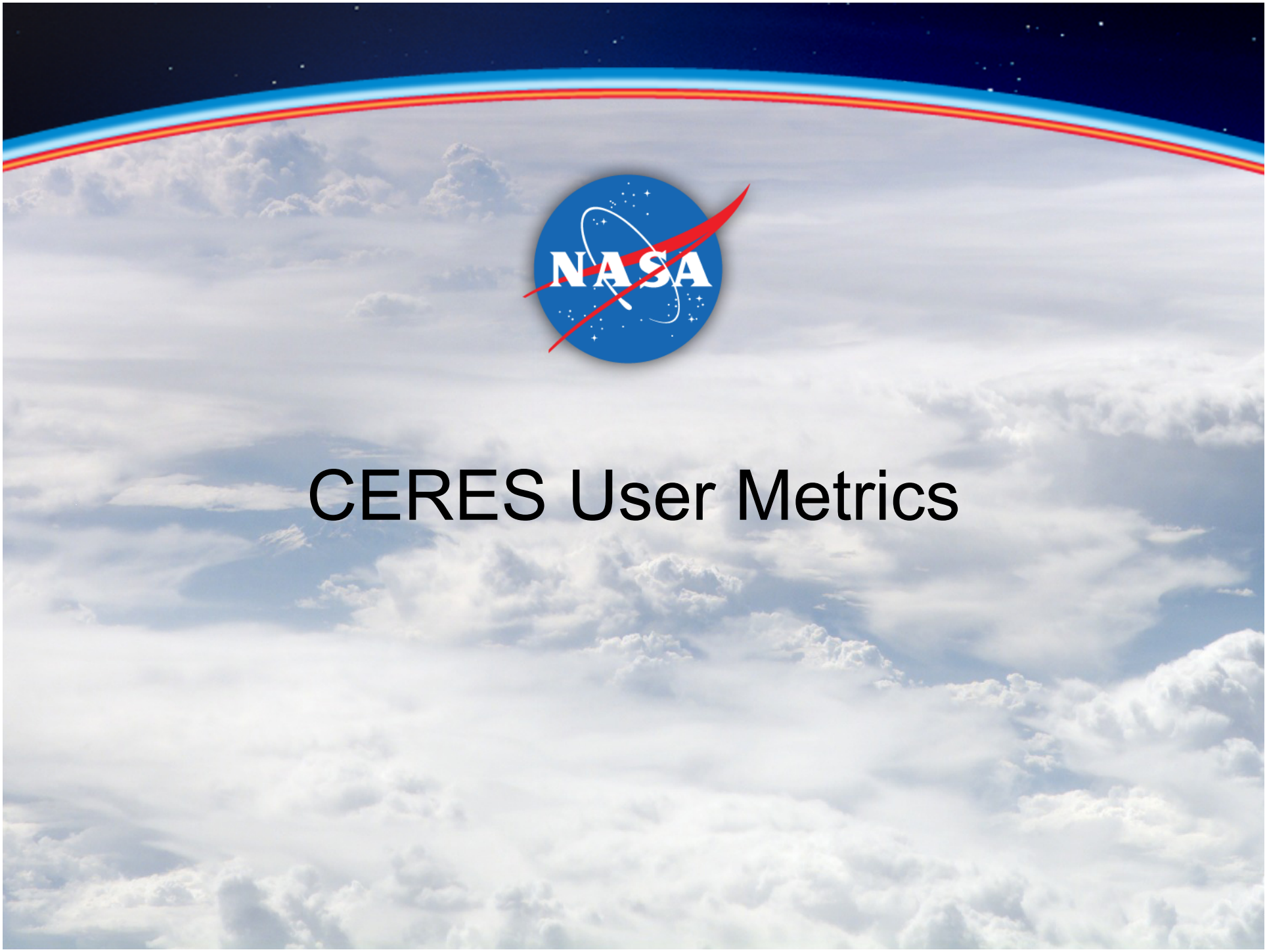
Atmospheric Science Data Center Update

CERES Science Team Meeting
October 4, 2011

Susan Sorlie, SSAI

Agenda

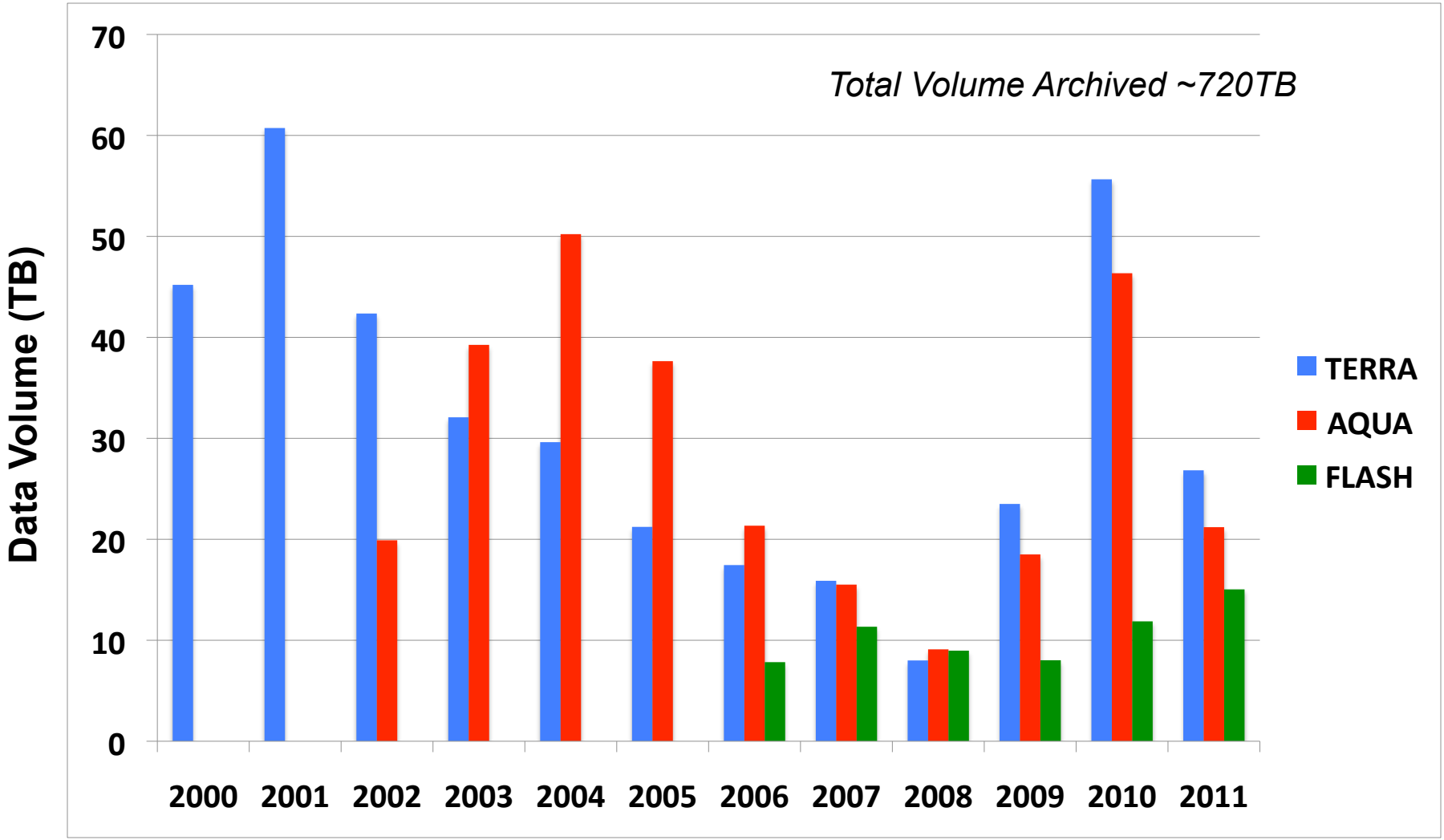
- CERES User Metrics
- NPP Readiness
- CERES/FLASHFlux Processing on AMI-P
- Data Access
- Conclusion



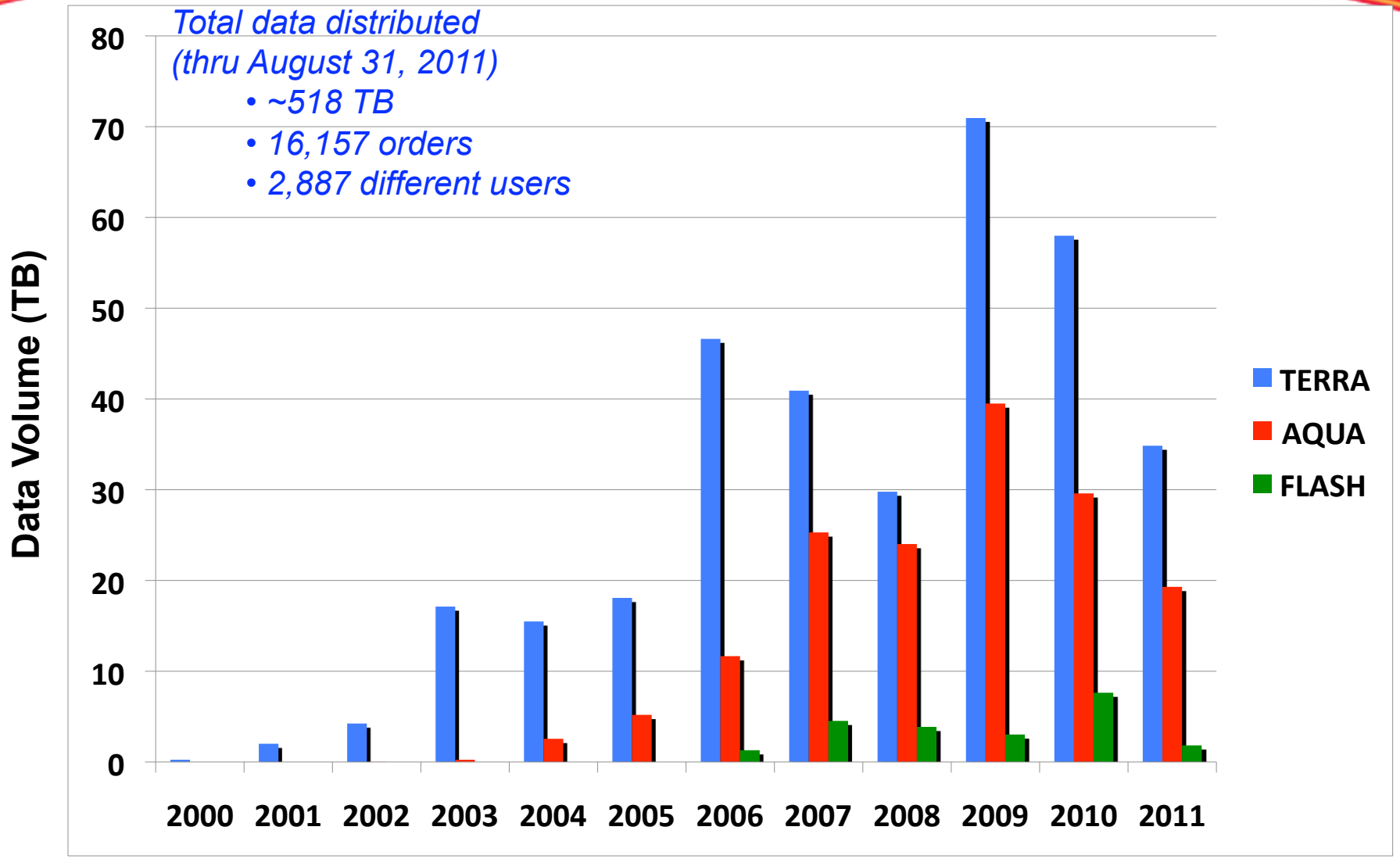
CERES User Metrics

CERES and FLASHFlux Archive Volume

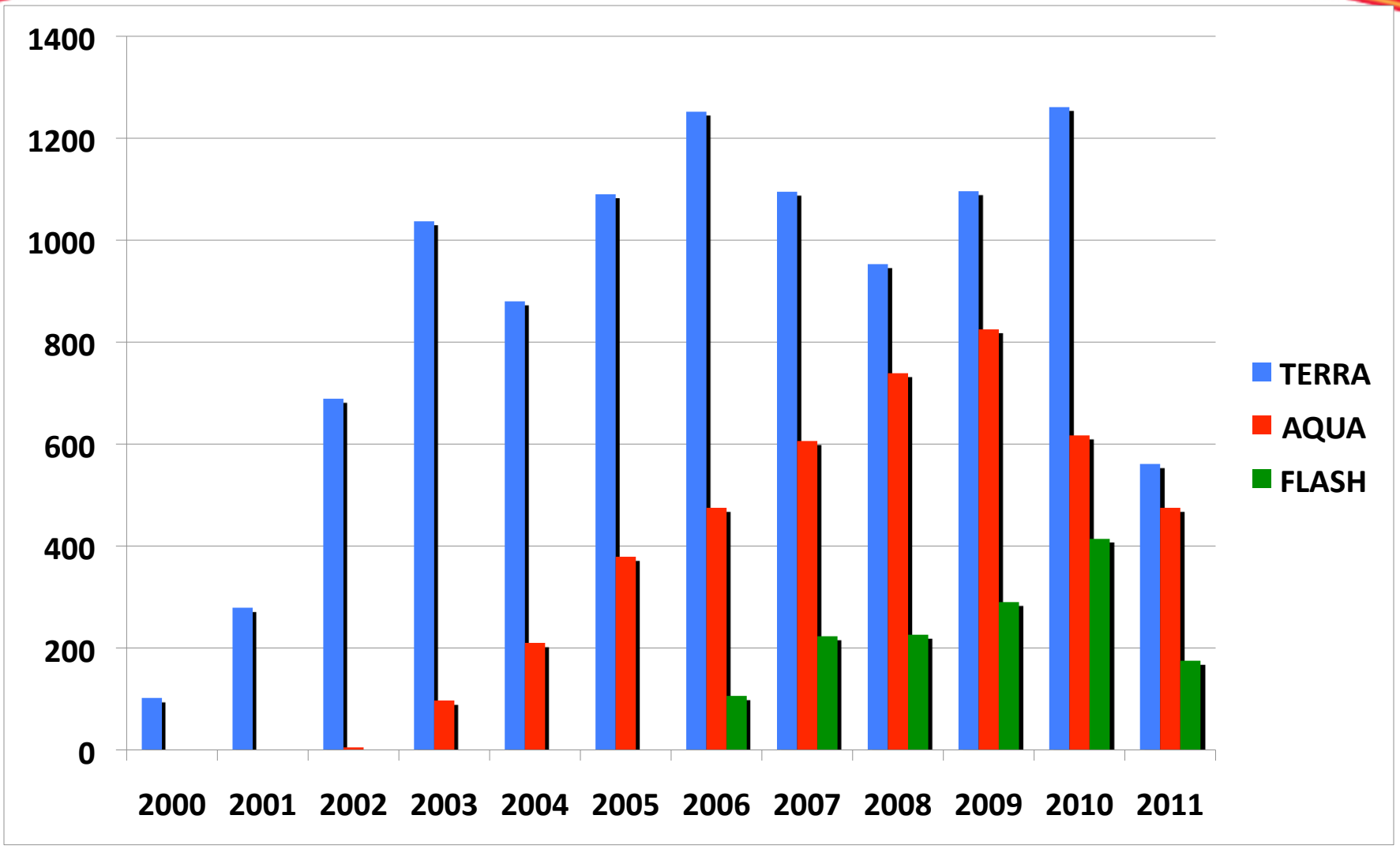
By Data Year



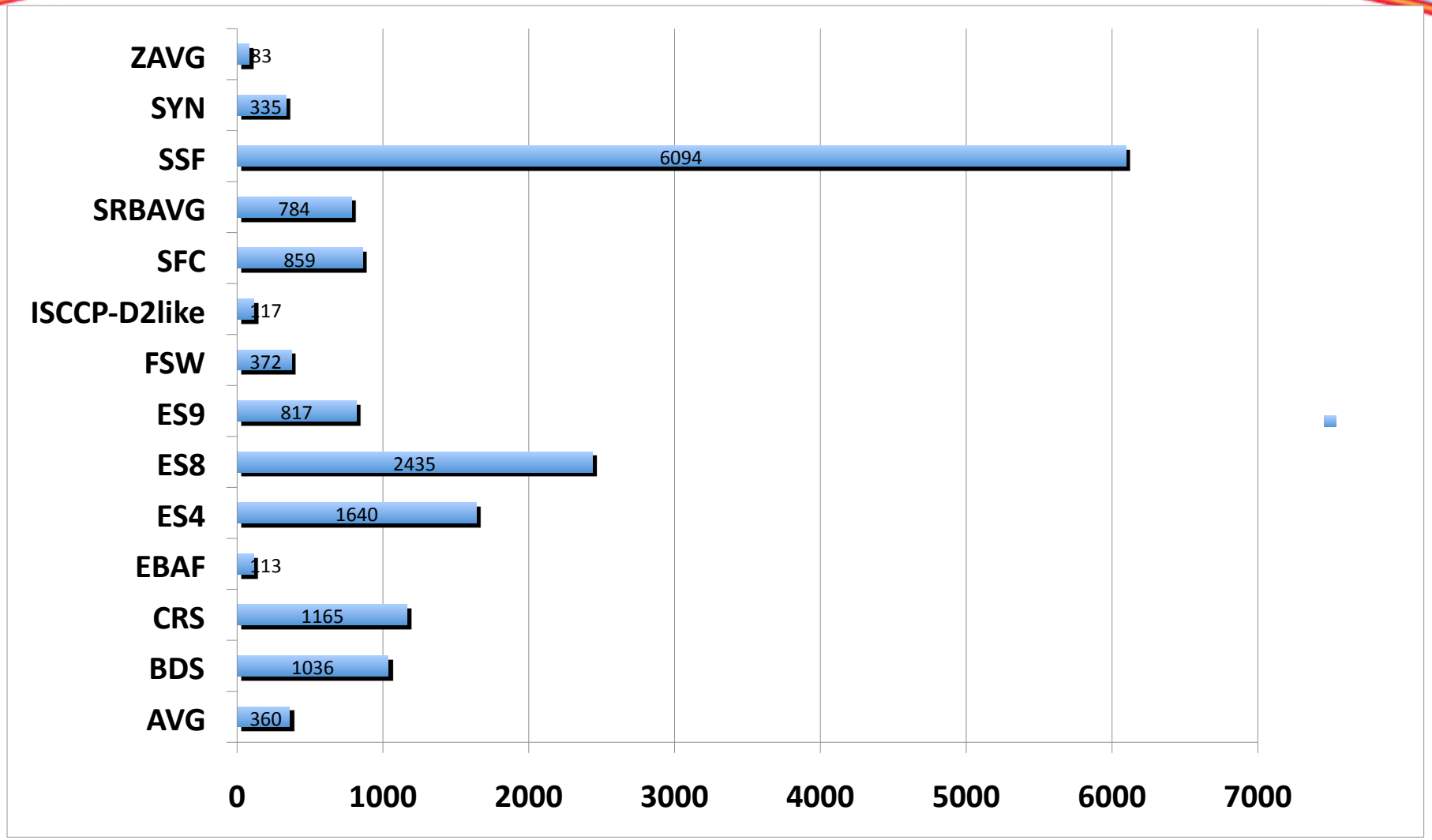
CERES and FLASHFlux Data Distribution



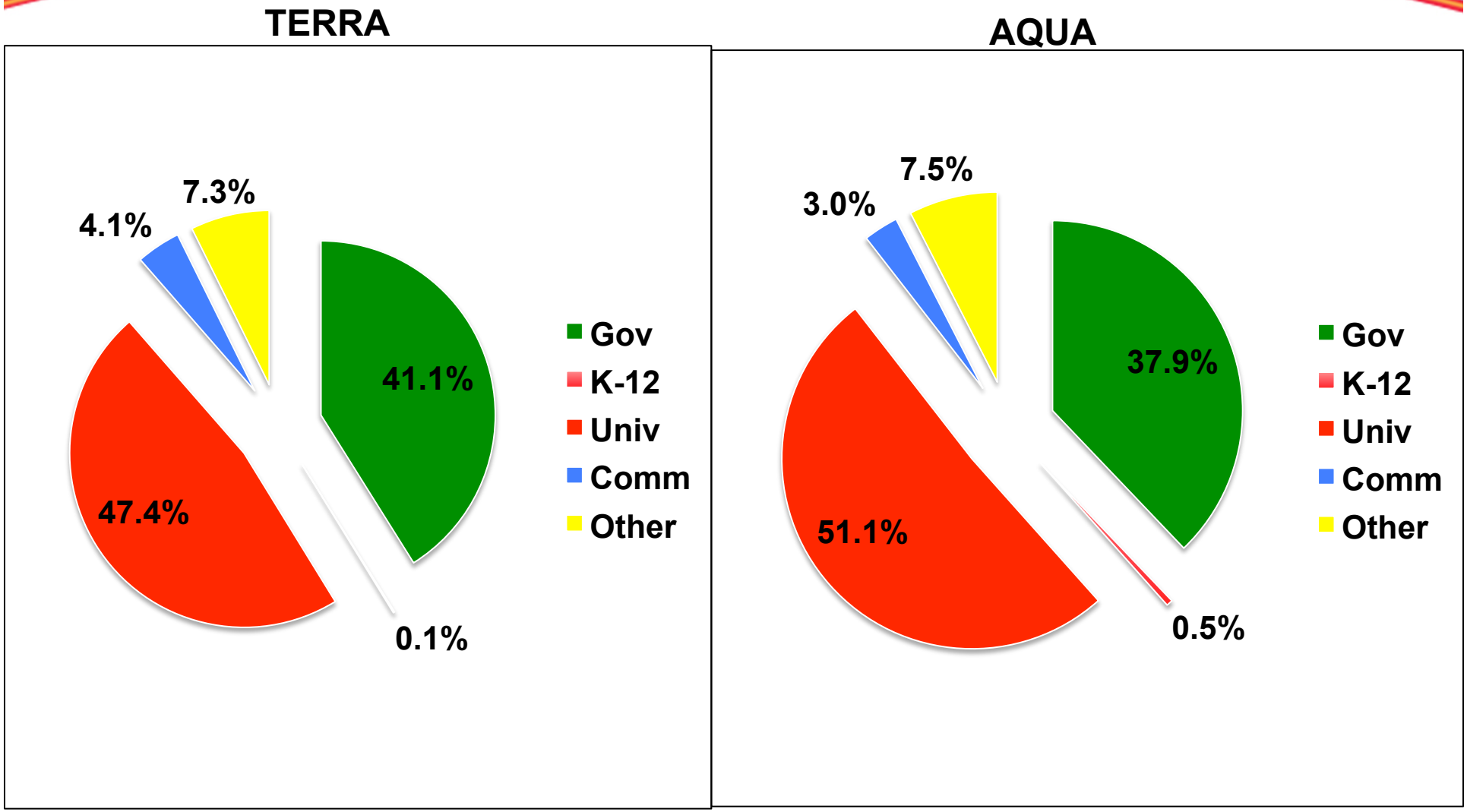
CERES and FLASHFlux Data Orders



CERES Orders by Product (Mar 2000- August 2011)

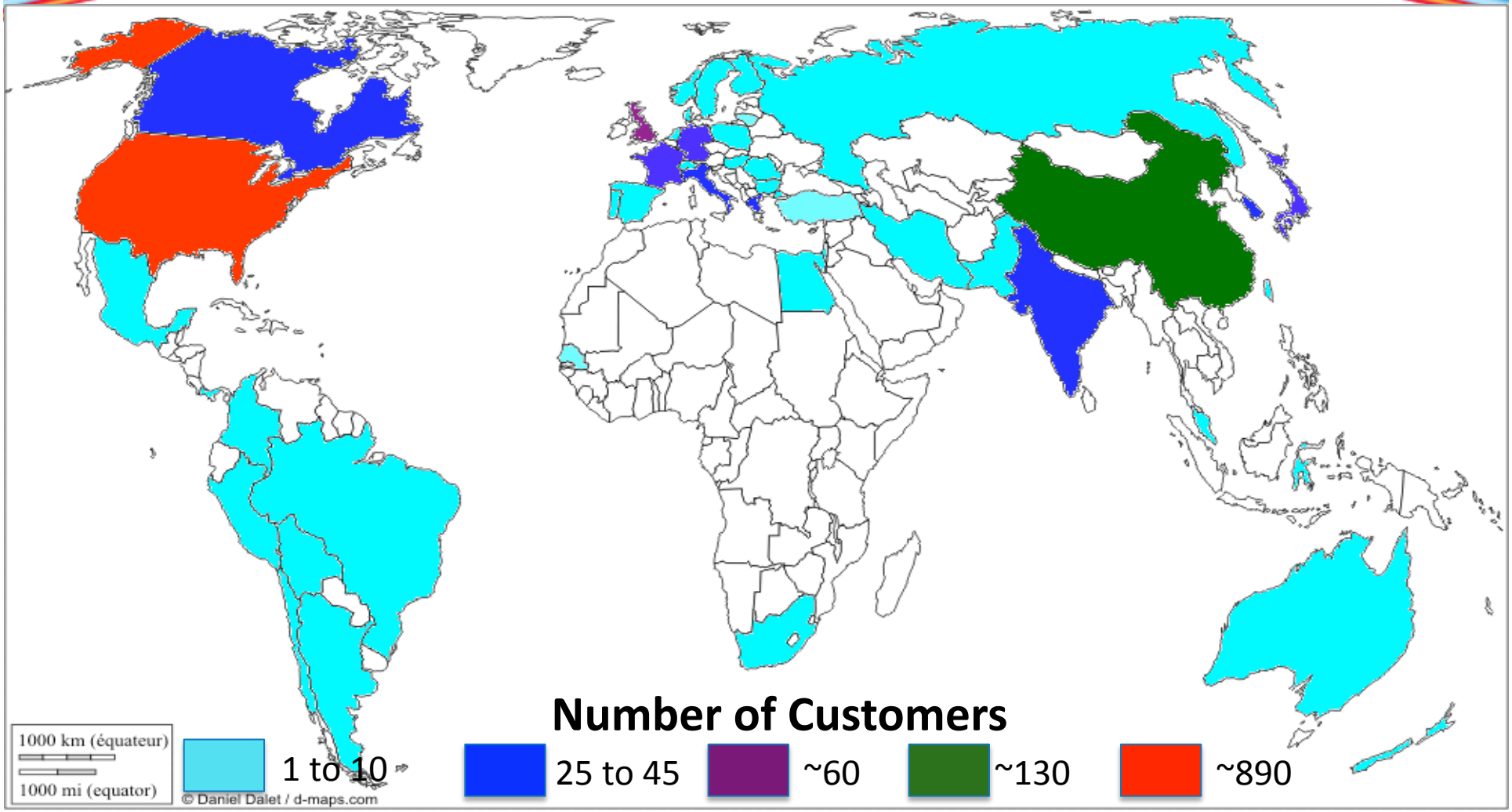


CERES Customers by Affiliation



ASDC CERES Data Users

Number of Customers



NPP Readiness

NPP Readiness

- ASDC is responsible for meeting applicable Level 3 NPP requirements
- All requirements were met (incrementally since the Ground System Interface Test in March 2010.)
- Meeting these Level 3 requirements constituted success during the NCT3 “LaRCStone”
- Participated in CERES NPP Science Operations Review on September 9
  - 4 RFAs: no significant concerns about ASDC readiness to support NPP

NPP Readiness

- The Atmospheric Science Data Center (ASDC) with the CERES Data Management Team (DMT) successfully completed the most recent run for record (NCT-4)
- All PGEs received were executed and completed with expected results
  - Four CERES Instrument Product Generation Executives (PGEs) were successfully run on the AMI-P system
  - One CERES ERBE-Like PGE was successfully run on the legacy (*magneto*) Linux system
- All products were archived and made available on the DPO for DMT analysis

NPP Readiness

- The RFR data flow was executed in addition to normal operations and had no impact on over-all system performance.
- Successful handling of the production loads and unexpected anomalies demonstrated the readiness of the ASDC, the operations team and environments for ingest, archive, distribution, and processing of FM5 data

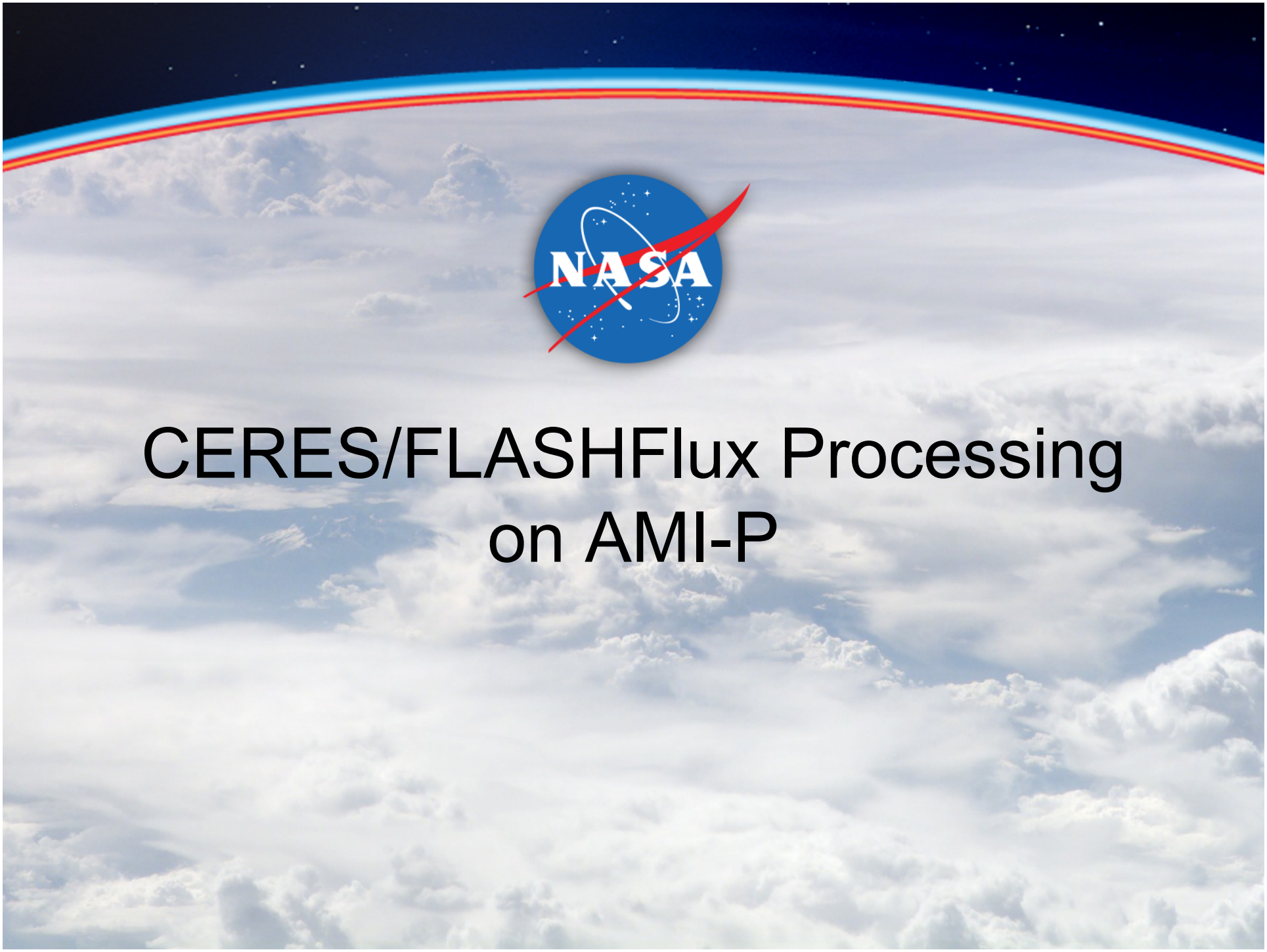
NPP Readiness

Handling Terra/Aqua/NPP

- ASDC resources will be allocated to work on highest priority requests

NPP Post Launch Cal-Val:

- ASDC ability to rapidly accommodate multiple code changes needed to support post launch operations requires
  - simple deliveries (deliver only what needs to be fixed)
  - accurate documentation
  - complete delivery of all components needed to run PGE
  - prioritization of resources



CERES/FLASHFlux Processing on AMI-P

CERES/FLASHFlux Processing on AMI-P

- AMI System was deployed to replace legacy computational resources (SGI, Sun Microsystems, Apple XSERVE/XSAN)
- AMI-P configuration is the *production* component of larger AMI system.
- Standard COTS software (SUSE Linux, GPFS, FORTRAN, C, C++, IDL, Grid Engine) and CERES custom items (PGEs, libraries, toolkits) are installed

CERES /FLASHFlux Processing on AMI-P

Compute Servers: 168 IBM Power 6 cores; 176 IBM x86 cores

- 42 Power 6 blades (168 P6 cores)
- 14 Intel Xeon Nehalem x86 blades (112 x86 cores)
- 2 x IBM 3950-M2 servers (32 x86 cores)
- IBM 3850-M3 servers (32 x86 cores)

Storage: IBM DS4800 Disk Systems

- Production workspace: 100 TB (IBM GPFS file system)
- AMI Data Products Online (DPO): 900 TB (IBM GPFS file system)
  - ASDC ANGe system populates DPO with input data required for production
  - Output data products made available to local science team for QA and analysis

CERES /FLASHFlux Processing on AMI-P

Ethernet Network: Cisco Nexus 7018

- 1 GigE VLAN: provides server-to-server communications
- 10GigE VLAN: provides compute servers access to GPFS file systems

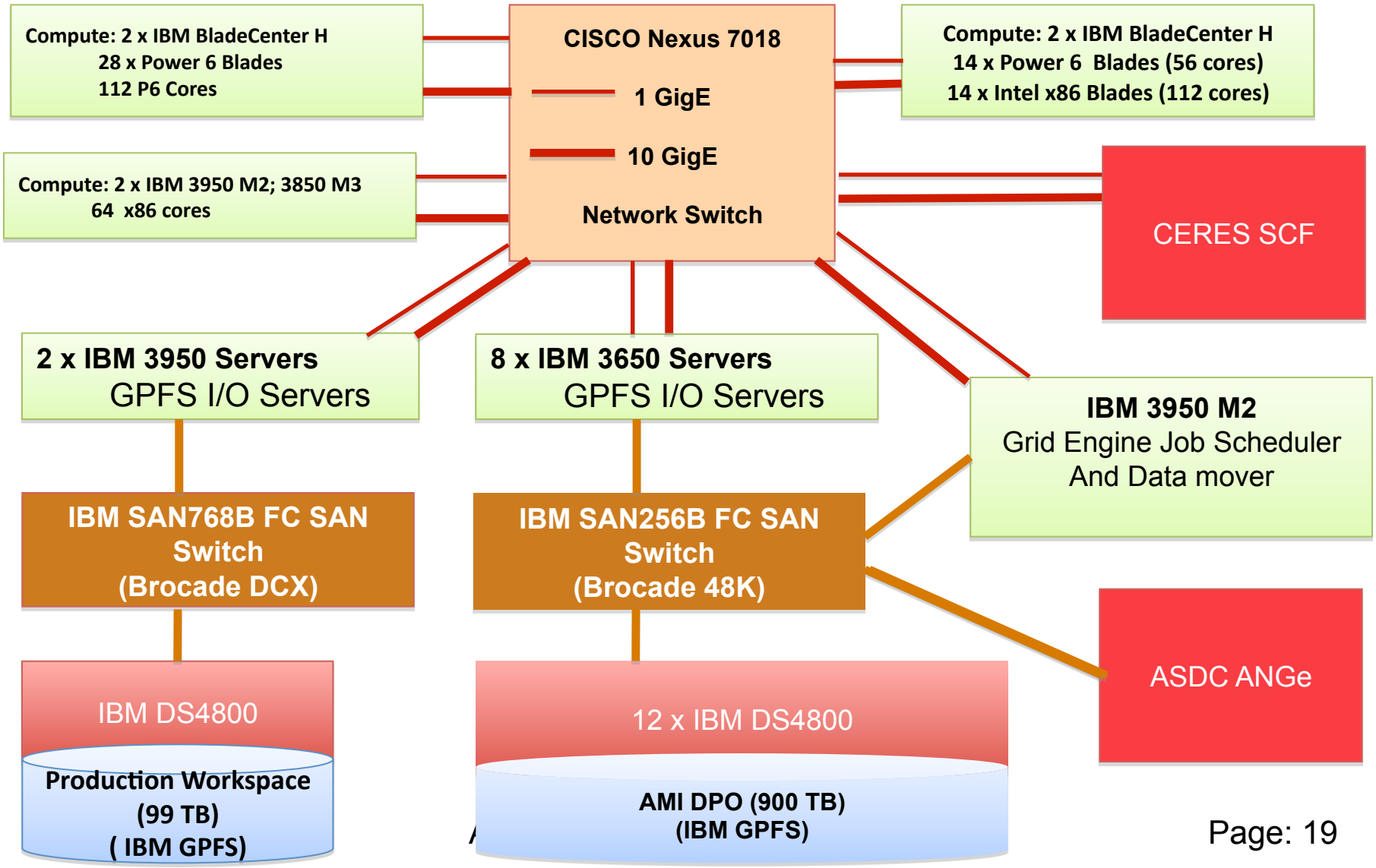
Storage Network: IBM SAN256B and SAN768B FC Switches

- Provides 4/8 Gbps fibre channel connections to I/O servers and disk storage

Hardware Maintenance:

- Compute and storage HW purchased in 2008 and 2009 with 5-year warranty
- Budget projections include funds required for hardware refresh or maintenance extensions (if more cost effective)

CERES /FLASHFlux Processing on AMI-P



CERES /FLASHFlux Processing on AMI-P

Present Status

- Instrument and MOA PGEs have been run in production
  - 15 PGEs promoted on AMI-P
    - 12 Instrument ; 1 TISA (SS8); 2 MOA
- ERBELike to be promoted this week
- Clouds delivered to ASDC (on hold)
- FLASHFlux being produced on both SGI (*warlock*) and AMI-P in parallel since August 1
- FLASHFlux is preparing to upgrade to GEOS 5.7.2; files from the Goddard DAAC are being delivered

CERES /FLASHFlux Processing on AMI-P

Near-Term Plans

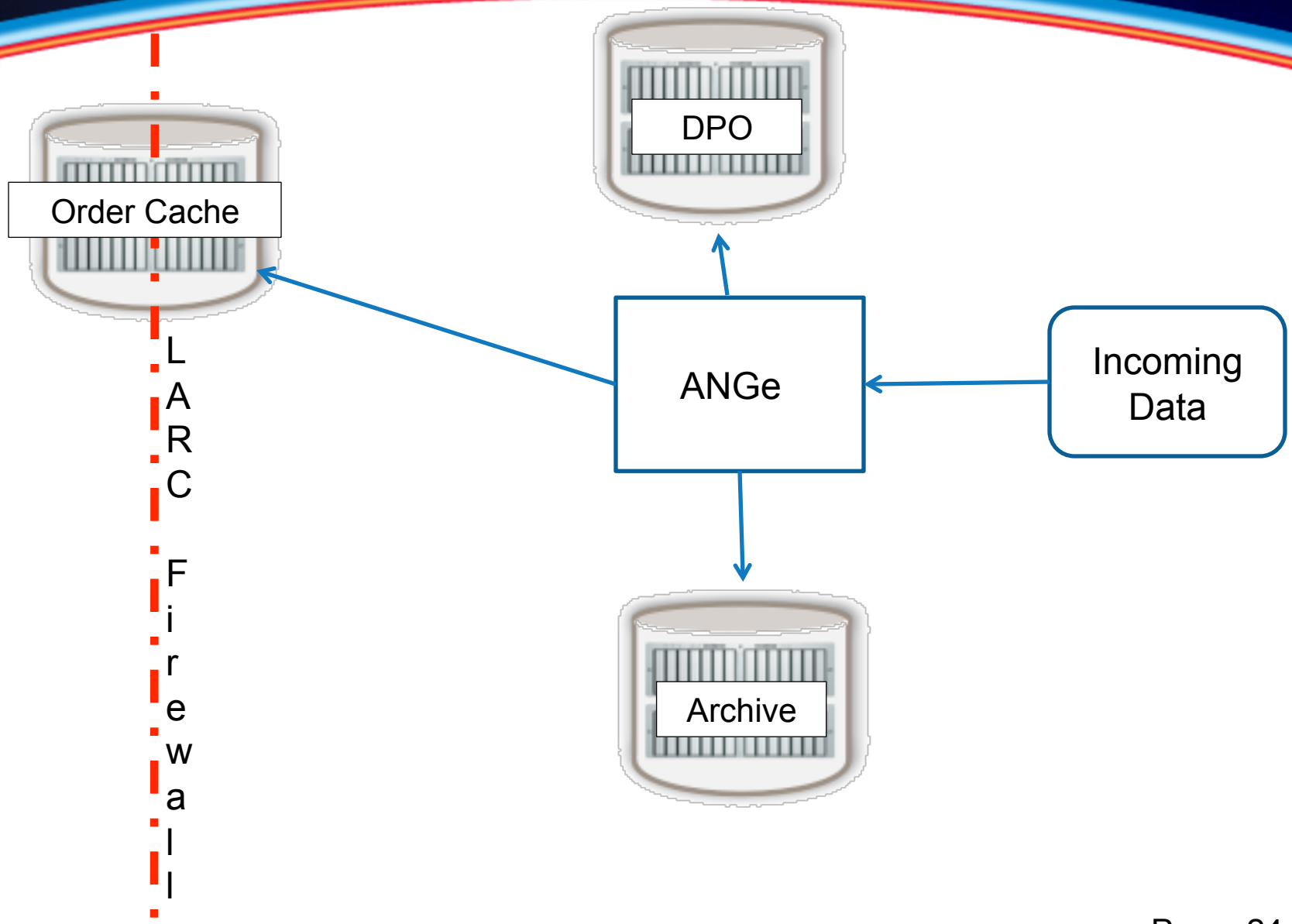
- Adding 56 P6 cores and 112 x86 cores
- Compute nodes access to DPO via 10 GigE
- Replace Open Source Sun (Oracle) Grid Engine with commercially supported version: Univa Grid Engine
  - Need to coordinate timing with CERES DMT
- Continue the migration of legacy web servers to new web architecture

Data Access

ANGe 2.0

- The promotion of ANGe 2.0 will encompass software and hardware changes that will improve back-end operations and utilize new hardware
  - Software changes will allow storage of files in multiple locations: new tape system; data products on-line (DPO) and order cache (a disk-based data-store accessible from outside LaRCNET)
  - Files will be stored by observation date
  - Storage configuration information will be maintained more efficiently
  - ANGe 2.0 will significantly improve efficiency and expand capability

ANGe 2.0: Store files to multiple locations



Subsetting

Approach

- Capability to subset data has been requested by the ASDC User Working Group and ASDC users
- ASDC responded with a concept that meets the following goals:
  - Provide a common back-end framework that would provide services for multiple products
  - Provide a flexible environment that handles product specific differences on a case by case basis
  - Will not force a “one size fits all solution” or complex design
  - Allow customers to provide their own user interface (UI) if they prefer to define the user’s experience based on customer relationship

Subsetting

Project	Products	Collaboration	Status
CALIPSO	L1 and L2 Lidar data	UI produced by project; maintained by ASDC	Beta-production; September 2010 release
CERES	L2 SSF data	UI produced and maintained by project	Beta-production; September 2011 release
TES	L2 and L3 data	UI produced and maintained by ASDC	Under development
MISR*	TBR	TBR	Future initiative

*A MISR subsetter interface already exists; this effort is to integrate with the subsetter framework

EOSWEB Re-design Effort

Goals

- Provide ASDC users with an “easier to use” interface that includes
  - Data Information
  - Data Ordering Mechanisms
  - Tools and Services
  - Access to external sites (i.e. ECHO, Giovanni)
- Improve the sustainability and maintainability by ASDC staff
- Modernize the site using current technologies
- Collaborate with stakeholders to ensure the needs of our user community are being met

EOSWEB Re-design Effort

Plan

- Kick-off Telecon – held September 14
- Survey – released September 15
- One-on-one discussions following the survey to clarify and elaborate responses
- Review of survey results
- Develop web prototypes
- Develop Requirements – January 2012
- Initiate Development

Conclusion

- Continual increase in CERES archive and distribution of products worldwide
- ASDC is ready to support NPP
- CERES and FLASHFlux production are making progress moving to AMI-P
- Subsetting efforts are progressing
- Effort to improve the user experience for those visiting ASDC data pages is underway